

UC Irvine

UC Irvine Previously Published Works

Title

Operational snow modeling: Addressing the challenges of an energy balance model for National Weather Service forecasts

Permalink

<https://escholarship.org/uc/item/52h5s118>

Journal

Journal of Hydrology, 360(1-4)

ISSN

0022-1694

Authors

Franz, KJ
Hogue, TS
Sorooshian, S

Publication Date

2008-10-15

DOI

10.1016/j.jhydrol.2008.07.013

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



available at www.sciencedirect.com



journal homepage: www.elsevier.com/locate/jhydrol



Operational snow modeling: Addressing the challenges of an energy balance model for National Weather Service forecasts

Kristie J. Franz ^{a,*}, Terri S. Hogue ^b, Soroosh Sorooshian ^a

^a CHRS-HSSoE, Department of Civil and Environmental Engineering, E-4130 Engineering Gateway, University of California, Irvine, CA, United States

^b Department of Civil and Environmental Engineering, 5731F Boelter Hall, University of California, Los Angeles, CA, United States

Received 31 August 2007; received in revised form 4 April 2008; accepted 10 July 2008

KEYWORDS

Streamflow forecasting;
Operational snow
modeling;
Coupled snow–runoff
model;
Hydrologic modeling

Summary Prediction of snowmelt has become a critical issue in much of the western United States given the increasing demand for water supply, changing snow cover patterns, and the subsequent requirement of optimal reservoir operation. The increasing importance of hydrologic predictions necessitates that traditional forecasting systems be re-evaluated periodically to assure continued evolution of the operational systems given scientific advancements in hydrology. The National Weather Service (NWS) SNOW17, a conceptually based model used for operational prediction of snowmelt, has been relatively unchanged for decades. In this study, the Snow–Atmosphere–Soil Transfer (SAST) model, which employs the energy balance method, is evaluated against the SNOW17 for the simulation of seasonal snowpack (both accumulation and melt) and basin discharge. We investigate model performance over a 13-year period using data from two basins within the Reynolds Creek Experimental Watershed located in southwestern Idaho. Both models are coupled to the NWS runoff model [Sacramento Soil Moisture Accounting model (SACCSMA)] to simulate basin streamflow. Results indicate that while in many years simulated snowpack and streamflow are similar between the two modeling systems, the SAST more often overestimates SWE during the spring due to a lack of mid-winter melt in the model. The SAST also had more rapid spring melt rates than the SNOW17, leading to larger errors in the timing and amount of discharge on average. In general, the simpler SNOW17 performed consistently well, and in several years, better than, the SAST model. Input requirements and related uncertainties, and to a lesser extent calibration, are likely to

* Corresponding author. Present address: Department of Geological and Atmospheric Sciences, Iowa State University, 3023 Agronomy Hall, Ames, IA 50011, United States. Tel.: +1 515 294 7454.

E-mail address: kfranz@iastate.edu (K.J. Franz).

be primary factors affecting the implementation of an energy balance model in operational streamflow prediction.

© 2008 Elsevier B.V. All rights reserved.

Introduction

There has been much written regarding the potential for improving snowmelt estimation and the subsequent streamflow forecasts through better estimates of initial states and the use of advanced models and data products (Perkins, 1988; Day et al., 1989; Swamy and Brovio, 1997; Marks et al., 1999; Carroll et al., 2001; Walter et al., 2005; Shamir and Georgakakos, 2006; Lehning et al., 2006). The National Weather Service (NWS), responsible for short- and long-term streamflow predictions across the United States, uses the SNOW17 model (Anderson, 1973) as part of their river forecast system. The SNOW17 is a continuous, conceptual model that simulates snow accumulation and ablation using only temperature and precipitation as inputs (NWS, 2004). Alternatively, snow models using the energy balance method simulate the physical processes that affect the thermal energy content of the snow pack, such as sensible, latent and ground heat fluxes, using multiple meteorological variables as input (Lang and Braun, 1990). Two reasons for using the SNOW17 in operational forecasting have been stated: (1) air temperature data are readily available throughout the US in real-time, and (2) previous tests conducted on experimental watersheds showed that the SNOW17 produced results "at least as good as" those from the energy-aerodynamic method (Anderson, 1973, 1976). More recent studies show that temperature-based snowmelt models and energy balance snowmelt models perform equally well under most conditions (Ohmura, 2001; Zappa et al., 2003).

The SNOW17 simulated maximum SWE with slightly less bias when compared in distributed mode to three land surface models which use the energy balance snow method (Mitchell et al., 2004). However, other studies have concluded that the SNOW17 was unable to properly capture snowmelt timing for a topographically complex basin in the Sierra-Nevada because the model does not consider the effect of shading on solar radiation reaching the basin; suggesting the model will be unable to predict timing of streamflow if melt variations between shaded and sunny regions are enhanced in a warmer climate (Lundquist and Flint, 2006). Runoff processes are highly sensitive to the influence of terrain on radiative processes (Lehning et al., 2006). Assuming continued climate variability and uncertain model forcing (i.e. data outside the observed record) an energy balance model may be a more prudent choice for modeling future snow conditions. With the continuing advancement of remote sensing capabilities and numerical weather models providing the potential for high resolution forcing, a snow model which uses the energy balance method may now be a more viable option than when the SNOW17 was deemed more appropriate for operations in 1973.

The NWS has established the Advanced Hydrologic Prediction System (AHPS) to modernize forecast services and improve hydrologic predictions through the incorporation of verified science from the climatological, meteorological,

and hydrological communities (McEnery et al., 2005). A recent National Research Council (NRC) panel identified a gap between state-of-the-art modeling capabilities and those used in AHPS and concludes that the NWS needs to incorporate advanced hydrologic science into their hydrologic models (NRC, 2006). Snow models are a central component of hydrologic forecasting systems during times when snow and/or snowmelt are the dominate influence on the regional streamflow. Given recent snowpack declines in the western United States and the uncertain impact on water resources (Mote, 2003; Stewart et al., 2004; Mote et al., 2005; Maurer, 2007), accurate prediction of spring snowmelt will become increasingly important as western populations grow, and demand more water, and operational agencies have to manage water under climate conditions outside of the historical record.

Several snow model comparisons studies have included models specifically developed for hydrological forecasting (Anderson, 1976; WMO, 1986; Brubaker et al., 1996; Essery et al., 1999; Etchevers et al., 2002; Mitchell et al., 2004; Lei et al., 2007), however, there have been no recent studies that examined alternate snow models within the framework of the NWS River Forecast System (NWSRFS) (i.e. coupling with NWS Sacramento Soil Moisture Accounting model (SAC-SMA) (Burnash et al., 1973)). Although the NWS continues to develop their forecast system, such as through the current effort to create an open architecture Community Hydrologic Prediction System (CHPS) (Schaake et al., 2006), some version of the current hydrologic forecast system will remain the standard for some time to come. To integrate scientific advancements into current operational hydrologic prediction, researchers must consider the existing methods and system.

The objectives of our work include the following: (1) to evaluate the performance of an energy balance snow model against the current NWS model (i.e. SNOW17), (2) to test a coupled energy balance snow model—SACSMA against the coupled SNOW17—SACSMA, and (3) to identify the benefits and challenges associated with implementing an energy balance snow model within the current NWS forecasting framework. Given the pressure to introduce advanced models into operational forecasting, this study is a first step in addressing the feasibility of incorporating a more complex snow model (assuming adequate forcing is available) into the NWSRFS. This study supports ongoing research to evaluate energy balance snow models for operational hydrologic forecasting by the NWS Office of Hydrologic Development (OHD) (Lei et al., 2007) and activities at the NWS National Operational Hydrologic Remote Sensing Center (NOHRSC). NOHRSC operates the SNOW Data Assimilation System (SNODAS), a 1 km gridded energy balance model and data assimilation system, to assist in development of snow products (Carroll et al., 2001). NOHRSC produces daily areal snow cover and SWE products for the conterminous United States, which are distributed to the RFCs for guidance on updating

the SNOW17 model states (<http://www.nohrsc.noaa.gov/nsa/>).

We investigate the SNOW17 and the Snow–Atmosphere–Soil Transfer (SAST) model (Jin et al., 1999a,b), and compare their performance for simulating snow season processes and basin discharge, when coupled with the SACSMA. Numerous energy balance models are available for simulating snowmelt processes; we chose the SAST because it was easily available from previous studies (Jin et al., 1999a,b) and had shown comparable performance to other models in inter-comparison studies (Jin et al., 1999a; Nijssen et al., 2003). Both models are evaluated for simulation of point snow water equivalent (SWE), basin average melt, and discharge. The models are applied to a set of nested watersheds within the Reynolds Creek Experimental Watershed (RCEW), Idaho (Slaughter et al., 2001), where long-term continuous data sets were available.

Methods

The two snow models are initially evaluated at the point-scale for simulation of SWE. The comparison then progresses to the watershed scale utilizing a set of nested watersheds within the RCEW for simulation of discharge when coupled to the SACSMA. The combination of the SNOW17 and SAST with the SACSMA will be referred as SNOW/SAC and SAST/SAC, respectively.

Although there are various spatial scales (point, plot, grid, etc.) at which to evaluate model performance, we contend that models must first be understood (and evaluated) using the best possible source of data and at scales which reduce uncertainty. Rigorous investigation at a high spatial resolution with reliable ground-based data can lead to a more thorough assessment, and subsequent understanding, of model (structure and parameter) error prior to applying models at larger scales with alternative or new data sources (Hogue et al., 2006a). In addition, field observations and measurements are still considered the benchmark for hydrological information and understanding (Kirchner, 2006) and were, therefore, the preferred data source to evaluate the models in this study and create a baseline for future studies.

Finding basins with long-term observations in the US of all the variables required was difficult, therefore we focused this study on the RCEW basin for which most of the required data for the energy balance model were available. Online searches, literature review, and contact with persons in the hydrologic field were conducted over several years in an attempt to find multiple study sites. From the literature it is clear that most snow modeling studies are conducted with only 1–3 years of data (Essery et al., 1999; Strasser et al., 2002; Fierz et al., 2003; Xue et al., 2003; Etchevers et al., 2004), or 5 years in the case of the Sleepers River Experimental Watershed in Danneville, Vermont (1969–1974) (Brubaker et al., 1996; Yang and Niu, 2003; Sun and Chern, 2005). Studies often rely on estimated data from sources outside their basin of interest or computed values, such as radiation variables (Bowling et al., 2003). Piecing data together from multiple climate and snow observation sites is complicated by missing data, mismatch in the record, large distances between observations, and a

lack of web-based documentation about archives. Using alternative data sources such as the National Center for Environmental Prediction (NCEP), North American Regional Reanalysis (NARR), remote sensing products, and climate model outputs may introduce additional and unquantified uncertainties, making it more difficult to separate model and data errors. Although the use of remote sensing to force land surface models is being met with some success (Crow et al., 2006), the uncertainties associated with using this approach may negate rigorous comparison of model performance.

Study sites and data

RCEW is located in the Owyhee Mountains of southwestern Idaho and is characterized as having a semi-arid climate (Fig. 1). Seventy-five percent of the annual precipitation occurs as snow in the upper elevations of the basin (Hanson, 2001). Hourly climate (Hanson et al., 2001), precipitation (Hanson, 2001), and snow data (Marks et al., 2001) for water years 1984–1996 were available online for the basin.

Point-scale snow model evaluations were conducted with data from a small (0.39 km² (Pierson et al., 2001)) headwater basin within RCEW called Reynolds Mountain East (hereafter referred to as the East basin). The overall relief of the East basin is minimal (2024–2139 m). SWE data (both snow pillow and snow survey) and precipitation were collected in the center of the East basin (Marks et al., 2001). The data collection site is located in a grove of aspen and fir trees at an elevation of 2061 m (Marks and Winstral, 2001). Basin-scale snow melt and observed discharge were evaluated for both the East and the Tollgate basins. The Tollgate basin is 54.44 km² in area, and the elevation change from the outlet (1398 m) to the highest point (2244 m) is 846 m (Pierson et al., 2001).

Air temperature, relative humidity, solar radiation, and wind collected on the western ridge of the East basin (Fig. 2) (Hanson et al., 2001; Slaughter et al., 2001) were used for snow model forcing. This is an exposed shelf at an elevation of 2097 m and is characterized by sagebrush (Marks and Winstral, 2001). Basin average meteorological forcing for Tollgate were computed for the mean basin elevation (1837 m) based on the lapse rate between the observation point discussed above and the next closest observation point located at 1652 m (Fig. 2) (Hanson et al., 2001). The NWS uses a similar process to compute mean areal temperature for their forecast basins (Anderson, 2002). Basin average precipitation was computed using the Thiessen polygon method with the two precipitation gages in the East basin and nine precipitation gauges in the Tollgate basin. Given the minimal relief in the East basin, the use of the Thiessen polygon method is reasonable. Although the Tollgate basin has more substantial relief, precipitation gages are well distributed, including five gauges in the highest elevations, providing good representation of precipitation throughout the watershed.

Incoming longwave radiation was not observed in RCEW during the time period studied, and was estimated using procedures from standard published methods (Crawford and Duchon, 1999; Kimball et al., 1982; Steiner, 2001; Dingman, 2002). Longwave radiation contributed from air and

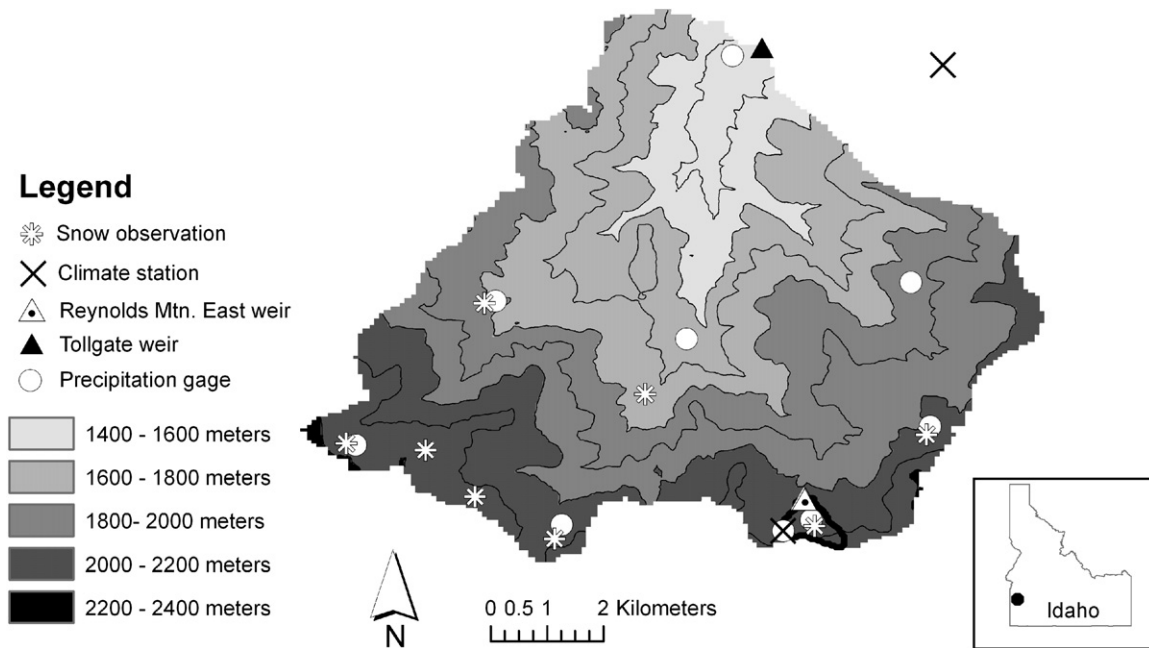


Figure 1 Location of the Reynolds Creek Experimental Watershed (insert) and the locations of observation points used in this study.

clouds was determined using the observations of solar radiation, dew point temperature, and vapor pressure (Franz, 2006). Climate data for the period September 1983–December 1984 were missing at the observation site at 1652 m. During this period the historical lapse rates for September to December from 1984 to 1996 were used to compute the values from the upper observation point to the mean basin elevation. The mean monthly temperature lapse rates were found to be negative for January 1985 and December 1985, but positive otherwise. The relative humidity was larger at the higher elevation site and varied $\pm 8\%$ between the two sites with the mean difference near $+2\%$ (Fig. 1). Wind speeds were greater at the upper site with a mean difference of 0.7160 m/s .

Models

NWS SNOW17

The SNOW17 model (Anderson, 1973) is based on an energy balance model described in Anderson (1968, 1976). Simplifications were made to the snowmelt calculations of the Anderson energy balance model to create the current version which has reasonable data requirements for operational applications. The SNOW17 (originally called HYDRO-17) is more complex than most traditional degree-day methods because the model has continuous accounting of the heat storage of the snowpack, as well as liquid water retention and transmission using empirically based relationships. However, the SNOW17 relies solely on temperature to model snowpack processes and, as a result, has traditionally been referred to as a temperature index model (Anderson, 1976).

Snow is modeled as a single layer in the SNOW17. The point-scale application of the SNOW17 requires eight

parameters (Table 1). When applied at the basin-scale, an areal depletion curve (ADC) and a parameter describing when less than 100% snow cover occurs (SI) are also required. The SNOW17 requires inputs of air temperature (T_a) and precipitation time series (Anderson, 2002).

Heat content in the snowpack increases or decreases as a function of the gradient between the antecedent temperature (as determined by the antecedent temperature index (TIPM)) and the current air temperature (Anderson, 1973). Melt occurs when enough heat has been added to the snowpack to bring its heat content to zero. During non-rain periods the depth of melt is determined by

$$M = M_f^*(T_a - \text{MBASE}) \quad (1)$$

where M is the depth of melt (mm), M_f is the seasonally varying melt factor ($\text{mm}/^\circ\text{C}$), and MBASE is the temperature above which melt will occur (typically set to 0°C). The melt factor is computed from a sinusoidal curve with limits defined by the maximum (MFMAX) and minimum (MFMIN) melt factor parameters (Anderson, 1973). Heat conduction through the snowpack is assumed to vary similarly to the non-rain melt factor and is scaled by a negative melt factor (NMF). Energy balance equations are used to compute melt during rain on snow events using several assumptions about meteorological conditions during rainy periods (Anderson, 1973).

Excess water will occur in the pack when it is isothermal at 0°C and the liquid water holding capacity of the pack (PLWHC) is met. Excess water is lagged and attenuated to simulate flow through the pack based on a series of empirically derived equations for ripe snow. A constant daily rate of melt at the soil–snow interface (DAYGM) is parameterized in the model to account for the geothermal heat flux at the ground surface (Anderson, 1973).

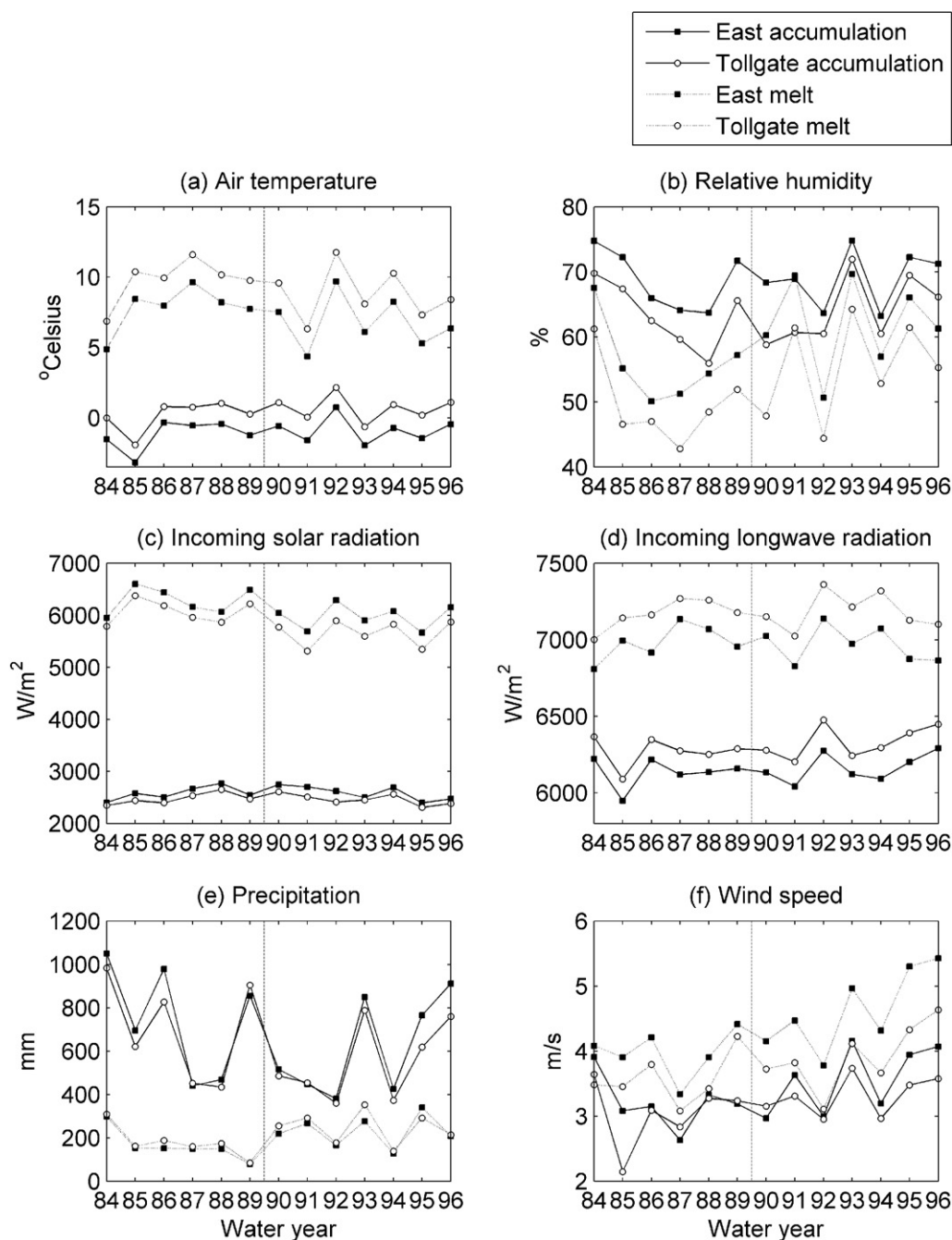


Figure 2 Daily average climate variables for the accumulation period (October through March) and melt period (April through June) at the East and Tollgate watersheds.

Snow–Atmosphere–Soil Transfer model (SAST)

The SAST model (Jin et al., 1999a,b) is based on the physical parameterizations of the SNTHERM model (Jordan, 1991) and Anderson (1976), but has been simplified to allow application for climate and hydrologic studies (Sun et al., 1999). Jin et al. (1999a) applied the SAST and the SNTHERM to simulate snow depth and SWE during the 1992 and 1993 melt seasons at the Mammoth Mountain in the eastern Sierra Nevada, California. The study concluded that the SAST, with the proper selection of layer depths, generated diurnal snowmelt characteristics similar to the SNTHERM and was less computationally expensive.

The SAST model accounts for heat conduction, snow compaction, grain growth, and melt. A maximum of three snow layers are used, which vary in thickness depending upon the total depth of snow (Sun et al., 1999). SAST computes the following state and output variables: SWE, snow density, melt, snow temperature profiles, heat content, and turbulent heat fluxes at the snow surface. Meteorological data required for the SAST model includes incoming and reflected shortwave radiation, incoming longwave radiation, air temperature, precipitation, wind speed, and relative humidity.

The dynamic albedo estimation method described by Dickinson et al. (1993) was applied. This method considers

Table 1 Description of the SAST and SNOW17 model parameters and the ranges used in the calibration

Parameter	Description	Range	Calibrated
SNOW17			
MFMAX	Maximum melt factor (mm/°C/6 h)	0.50–2.00	0.85
MFMIN	Minimum melt factor (mm/°C/6 h)	0.05–0.90	0.05
PLWHC	Liquid water holding capacity (%)	0.02–0.3	0.01
MBASE	Melt base temperature (°C)	0.0–1.0	0.0 ^a
NMF	Maximum negative melt factor (mm/°C/6 h)	0.05–0.50	0.15 ^a
DAYGM	Average daily ground melt (mm/day)	0.0–0.3	0.1 ^a
UADJ	Wind function factor (mm/mb)	0.03–0.19	0.15 ^a
TIMP	Antecedent snow temperature index	0.1–1.0	0.4 ^a
Parameter	Description	Original	Calibrated
SAST			
R3	Albedo decay parameter	0.3	0.4
BEXT	Near Infrared extinction coefficient	400	320
CV	Visible extinction coefficient	0.003036	0.003000
FLMIN	Minimum liquid holding water capacity (%)	0.03–0.10	0.01
ZNAUGHT	Roughness length	0.001–0.002	0.001
DZMAX	Maximum thickness of surface snow layer (m)	0.02	0.01
DZMIN	Minimum thickness of surface layer (m)	0.010	0.005
DZNMAY	Maximum thickness of second snow layer (m)	0.20	0.15
DZNMIN	Minimum thickness of second snow layer (m)	0.18	0.10
AVO	New snow albedo	0.95	0.90

^a Parameter values transferred without adjustment from the North West River Forecast Center's Upper Owyhee basin.

albedo for the near infrared and visible ranges, requires no data in addition to what was available, and showed comparable results for SWE simulation when tested against other methods for this study. The albedo averaging method of Tarboton and Luce (1996) was applied to snow depths less than 0.1 m to account for the changing surface albedo as the snowpack becomes shallow. The SNOW17 areal depletion curve (ADC) parameterization was adapted for use with the SAST model and, similarly, an SI parameter was set to initialize use of the ADC when there was less than 100% snow cover.

The SACramento Soil Moisture Accounting model (SACSMA)

The SACSMA is the rainfall–runoff model used by the NWS as part of their streamflow forecasting system. The SACSMA is a saturation excess model which represents percolation, soil moisture storage, drainage, and evapotranspiration (ET) processes in a conceptual manner (Burnash et al., 1973). Input to the SACSMA is precipitation and/or snowmelt. Potential evaporation (PE) values for the 16th of each month are used to linearly interpolate daily PE. PE values from an adjacent NWS forecast basin (upper Owyhee River basin) were obtained from the North West River Forecast Center (NWRFC) and used directly for the RCEW. The SACSMA has 16 parameters, four of which are typically set to default values (Table 2). Model output is basin average runoff depth.

Model calibration

Both snow models were manually calibrated using the 13-year record of SWE observations at the East snow pillow site. Model calibrations and evaluations were conducted

according to water years to cover an entire snow season from accumulation through ablation. To simplify the presentation of results, years are named by the water year represented, e.g. 1993 represents the water year October 1, 1992–September 30, 1993.

The SAST was first assessed using default parameter values. However, the default parameters resulted in a general overestimation of SWE at the East snow pillow site and significantly higher SWE values in water years 1993, 1995, and 1996. Because there was no specific guidance regarding identification of SAST parameters or proper parameter ranges, a set of possible variables were identified and adjusted using Xue et al. (2003), Sun et al. (1999), Jordan (1991), and Anderson (1976) as guidance (Table 1).

The SNOW17 model calibration was guided by parameter ranges and suggested values given in Anderson (2002), NWSRFS documentation, and parameters obtained from the NWRFC forecast basin. The rain/snow threshold value was set to 1 °C for each model (the value used by NWRFC in the adjacent headwater basin), assuring both models would receive equal amounts of snowfall input. There were no corrections made to the mass of precipitation input.

SACSMA parameters are conceptual in nature and are sensitive to spatial and temporal scales as well as inputs (Finnerty et al., 1997), therefore we did not assume that the same SACSMA parameters could be used with both snow models or in both watersheds. Although other studies have successfully used a priori parameters for a conceptual runoff model coupled with a snow model (Lehning et al., 2006), use of parameters from an a priori estimation method for SACSMA (Koren et al., 2000) were found to be unsatisfactory for these study basins. The SACSMA model parameters from the adjacent NWRFC forecast basin were tested with poor results. Initial

Table 2 SACSMA parameters and calibrated values for the SAST (SAST/SAC) and SNOW17 (SNOW/SAC) model outputs

Parameters	Description	Range	Reynolds Mountain East		Tollgate	
			SAST/SAC	SNOW/SAC	SAST/SAC	SNOW/SAC
UZWIM	Upper-zone tension water maximum storage (mm)	1–150	1	15	3	13
UZFWM	Upper-zone free water maximum storage (mm)	1–150	36	12	24	72
UZK	Upper-zone free water lateral depletion rate (day^{-1})	0.100–0.800	0.102	0.100	0.417	0.319
PCTIM	Impervious fraction of the watershed (decimal fraction)	0.00–0.10	0.01	0.03	0.01	0.00
ADIMP	Additional impervious area (decimal fraction)	0.00–0.40	0.02	0.13	0.08	0.09
ZPERC	Maximum percolation rate (dimensionless)	1.0–250.0	13.8	8.7	188.2	139.2
REXP	Exponent of the percolation rate (dimensionless)	0.5–5.0	1.2	0.5	2.1	1.3
LZWIM	Lower zone tension water maximum storage (mm)	1–500	168	80	186	148
LZFSM	Lower zone free water supplementary storage (mm)	1–1000	8	2	1	2
LZFPM	Lower zone free water primary storage (mm)	1–1000	130	160	951	253
LZSK	Lower zone supplementary free water depletion rate (day^{-1})	0.010–0.500	0.500	0.495	0.479	0.423
LZPK	Lower zone primary free water depletion rate (day^{-1})	0.0001–0.2000	0.002	0.002	0.002	0.002
PFREE	Fraction of water percolating from upper zone directly to lower zone free water storage (decimal fraction)	0.01–0.80	0.23	0.25	0.12	0.12
KROUTE	reservoir routing exponent	0.10–1.00	0.10	0.15	0.21	0.42
<i>Fixed parameters</i>						
RIVA	Riparian vegetation area (decimal fraction)	—	0.00	0.00	0.00	0.00
RSERV	Fraction of lower zone free water not transferrable to lower zone tension water (decimal fraction)	—	0.30	0.30	0.30	0.30
SIDE	Ratio of deep recharge to channel baseflow (decimal fraction)	—	0.00	0.00	0.00	0.00

transfer of the SACSMA parameters from the smaller East basin to the larger Tollgate basin resulted in overestimation of peak flow and negative Nash–Sutcliffe efficiency values (NSE). Therefore, the SACSMA was calibrated for each snow model output in each basin using the Multi-step Automatic Calibration Scheme (MACS) (Hogue et al., 2000, 2006b). In this procedure, the Shuffle Complex Evolution (SCE) (Duan et al., 1992) optimization algorithm is run in three consecutive steps. In each step, the optimized parameters and the objective function are changed to focus the calibration on different parts of the hydrograph. Automatic calibration was chosen to identify SACSMA parameters because automatic calibration techniques for the SACSMA parameters have proven successful (Duan et al., 1992; Sorooshian et al., 1993; Yapo et al., 1996; Gupta et al., 1999; Hogue et al., 2000, 2006b), and multiple parameter interactions make manual calibration difficult for the SACSMA. In this study, the SACSMA is used as a transformation tool to determine if there is a realistic relationship between the snow model outputs and the streamflow, and to determine compatibility between the SAST and SACSMA.

Parameter ranges used in the optimization procedure were taken from previous studies (Boyle et al., 2000; Hogue et al., 2000, 2006b) and known RFC values, and were set wide enough to assure that the SACSMA parameter space was adequately sampled (Table 2). Runoff was routed to the basin outlet using a series of linear reservoirs defined by a single parameter (KROUTE). After preliminary testing, a single reservoir for the East and a series of five reservoirs for the Tollgate were found to provide adequate lag.

In the absence of spatial data, an ADC used by the NWRFC was chosen for the East watershed simulations. The ADC for the Tollgate was constructed by plotting estimated snow covered area (SCA) versus the basin average SWE divided by the maximum SWE for the year. Basin average SWE and SCA were found using Thiessen polygon weighted snow survey data, a reasonable method for determining SCA (Lang, 1986). The computed ADC for Tollgate follows the recommendation by Anderson (2002) that mountainous regions tend to have a combination of what he termed type B and C curves. The SI parameter was estimated from the maximum SWE values for Tollgate as per the SNOW17 model

guidance and verified with the discharge simulations. In the East, this initial estimate was found to be unsuitable; therefore, SI was adjusted during the SACSMA calibration procedure and by comparing the timing of the rain–melt output to the observed discharge.

Model evaluation

All model simulations were made at a 1-h timestep and aggregated for evaluation at the daily timestep. The model simulations were evaluated at the daily timestep using mean error (ME), root mean squared error (RMSE), Nash–Sutcliffe efficiency measure (NSE), percent bias (Pbias), and correlation coefficient (R):

$$ME = \frac{1}{N} \sum_{t=1}^N x_t - y_t, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - y_t)^2}, \quad (3)$$

$$NSE = 1 - \left(\frac{\sum_{t=1}^N (x_t - y_t)^2}{\sum_{t=1}^N (y_t - \bar{y})^2} \right) \quad (4)$$

$$Pbias = \left[\frac{\sum_{t=1}^N (x_t - y_t)}{\sum_{t=1}^N (y_t)} \right] \times 100 \quad (5)$$

$$R = \frac{n \sum_{t=1}^N x_t \times x_t - \left(\sum_{t=1}^N y_t \right) \times \left(\sum_{t=1}^N x_t \right)}{\sqrt{n \sum_{t=1}^N y_t^2 - \left(\sum_{t=1}^N y_t \right)^2} \times \sqrt{n \sum_{t=1}^N x_t^2 - \left(\sum_{t=1}^N x_t \right)^2}} \quad (6)$$

where x is the model output at time t , and y is the observation at time t . The timing errors of complete snowpack melt (melt-out error) and the peak discharge (peak discharge timing error) are computed where a positive (negative) value indicates the model simulated the variable later (earlier) than the observation occurred. The snow duration error is the difference in the number of simulated minus observed snow cover days throughout the water year.

Results

Point-scale snow model comparisons

SNOW17 parameter values for MFMIN, MBASE, NMF, DAYGM, UADJ, and TIMP were not changed from the NWRFC values because the simulation was found to be relatively insensitive to these parameters (Table 1). MFMAX and PLWHC were changed by +30% and –20%, respectively, from NWRFC values. Calibration of MFMAX resulted in the greatest improvements in both snow accumulation and melt. Adjusting the PLWHC improved the errors in SWE over-estimation, but had a lesser impact on melt timing.

SAST simulations were most sensitive to changes in the allowable maximum and minimum thickness of the snow layers (DZMAX, DZMIN, DZNMAX, DZNMN) (Table 1). Reducing the new snow albedo (AVO) and the minimum liquid water holding capacity (FLMIN) had a less significant impact on over-accumulation of snow in the SAST. Calibration of four other parameters (R3, BEXT, CV, and ZNAUGHT) resulted in only minor improvements by comparison. Adjusting SAST

parameters caused melting to begin earlier in the season; extended the melting period; and reduced the average magnitude of daily melt. After calibration the mean daily percent bias and mean daily error were reduced from 43% to 13% and from 107 to 34 mm, respectively.

Compared to the SNOW17, the SAST tended to have a larger overestimate of SWE, to begin melt at a later date, and to melt more quickly in the spring (Figs. 3 and 4 and Table 3). Summary statistics compare the SWE simulations to SWE observations from the snow pillow. Snow pillow observations have better correlation with snow core measurements during the accumulation period; errors during the melting period are variable and during repeated freeze thaw cycles can become uncertain (Sorteberg et al., 2001). At Reynolds Creek Watershed, the snow pillow observations tend to be underestimated compared to the snow survey during the wettest years (Marks et al., 2001). Therefore, there is greater uncertainty in the accuracy of the snow pillow data during the March through May period for the wettest years in this basin. The SAST matches the snow survey data slight better from January through March 1984, but in the remaining years the tendency by the SAST to overestimate SWE is validated by both the snow pillow and snow survey observations (Fig. 3).

The mean daily Pbias of the SNOW17 was 2.6% lower than the SAST and the average peak SWE error was 21.7 mm lower (Table 3). The SNOW17 had a higher NSE (0.95) as compared to the SAST (0.84) over the 13-year period. From 1984 to 1991, and during 1994, both model simulations resulted in fairly low mean daily SWE errors (9-year average of 9.6 mm for the SNOW17 and 4.6 mm for the SAST) and high NSE values (9-year NSE of 0.98 for the SNOW17 and 0.94 for the SAST). In 9 out of the 13 years, the SAST had a higher peak SWE error and a lower correlation than the SNOW17, although there is often minimal difference in simulated daily SWE as illustrated by 1984 and 1988 (Figs. 3a, b and 4). The largest errors for both models and greatest differences in performance occurred in 1992, 1993, 1995, and 1996. The average Pbias for these four years was 42% for the SAST, compared to 24% for the SNOW17. Efficiency values (NSE) for these same years were 0.87 for the SNOW17 and 0.63 for the SAST.

Figs. 3 and 4 illustrate performance of the two models in two drier years (1988 and 1992) and two wetter years (1984 and 1995) at the East site. As illustrated by the very dry conditions in 1992 and very wet conditions in 1995 (Figs. 3c, d and 4) the increased errors are not associated with either very large or very small snowpack. Both models showed fairly high accuracy during 1984 (wet) and 1988 (dry). Years 1993, 1995, and 1996 were characterized by above average precipitation (Fig. 2e); however, large accumulation does not appear to be a distinguishing reason for the poorer model performances. The models show fairly good performance (low Pbias and high NSE) during 1984 (Fig. 4b) which had the largest snowpack for the period of record.

There is little correlation between average daily longwave values in the accumulation periods and the tendency of the SAST to over-accumulate SWE. 1984 and 1995 had very similar average longwave inputs (Fig. 2), but SWE is overestimated in 1995 and not in 1984 (Fig. 3). In addition, 1992 had above average values of longwave radiation and also displayed over-accumulated SWE. Winter solar radiation inputs varied little from year to year, and no related

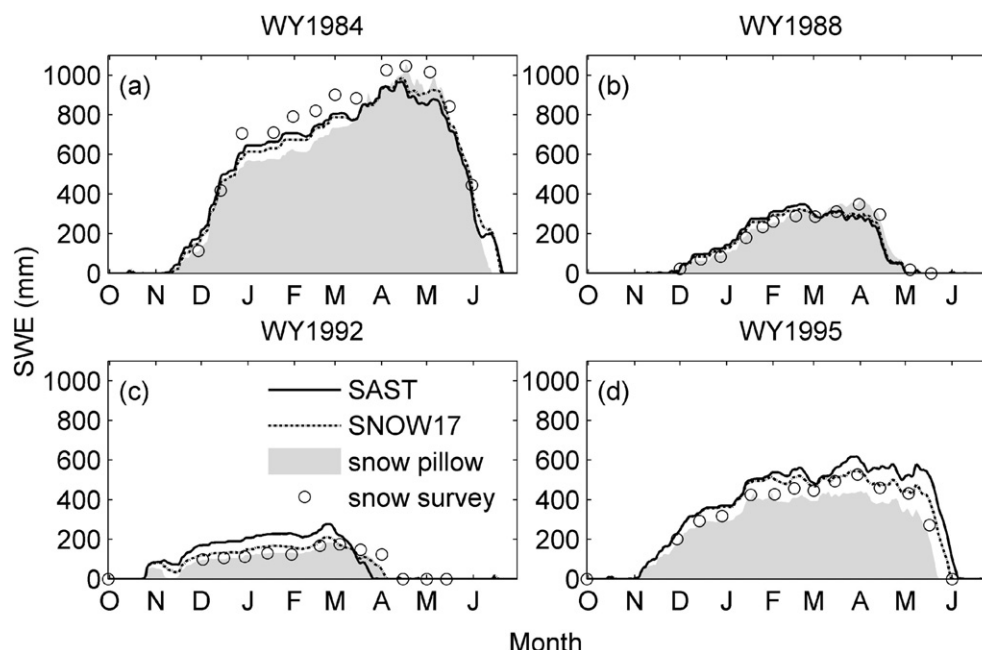


Figure 3 Daily snow water equivalent for the East basin study site for water years 1984, 1988, 1992, and 1995. Hourly observed SWE from the East basin snow pillow is depicted as the shaded region, and observed SWE from the snow survey is shown as the open circle. Selected years are shown for two wet (WY 1984 and WY 1995) and two dry (WY 1988 and WY 1992) seasons.

trend in SAST accumulation is observed. Longwave radiation was above normal throughout the 1992 snow season and the SAST model displays rapid melt around the middle of March 1992. Both 1984 and 1995 melt periods have below normal longwave radiation and solar radiation inputs, this correlates to a later melt in the SAST compared to the SNOW17 (Figs. 2, 3a, d and 4c).

The sensitivity of the model simulations to individual input data errors was investigated further by adding positive and negative biases of magnitudes 5%, 15%, and 25% to the model inputs. A continuous 13-year model run was generated using each altered data set. Temperature biases resulted in a slightly larger change in model performance for the SNOW17 compared to the SAST, indicating the SNOW17 has a higher degree of sensitivity to temperature at this site (Fig. 5). The SAST is least sensitive to biases in wind speed and most sensitive to biases in the radiation inputs. On average, a longwave radiation bias of +5 to +15% would improve the timing of the SAST melt in the spring and slightly improve the over-accumulation of snow beginning around February (Fig. 6). The SAST is most sensitive to biased data during the melt period and neither solar radiation nor longwave radiation errors significantly affected the accumulated snow during the late fall and early winter. Simulation of complete snowpack melt ranged from mid-May into July when $\pm 25\%$ longwave biases were introduced to the inputs (Fig. 6c). Comparably, the same bias in solar radiation had a lesser impact on SWE accumulation and melt.

Watershed scale snow model comparisons: coupling to the SACSMA

Reynolds Mountain East (East) watershed

Fig. 7a, c, e, and g illustrates the cumulative distribution of simulated melt from the two models. The SNOW17 produces

a minimum daily melt output throughout the snow covered period (due to the DAYGM parameter), whereas the SAST often has little to no melt through the winter and into February and March. This lack of melt early in the season contributes to the larger accumulation errors in the SAST, illustrated by the 13-year average mean peak SWE errors: 26 and 4 mm for the SAST and SNOW17, respectively (Table 3). The SNOW17, on the other hand, typically has more accurate SWE values going into the spring melt season (Fig. 3g and d).

The late melt and overestimation of SWE by the SAST leads to erroneous simulated peak streamflows in the spring and negative NSE values (illustrated during 1992 and 1995; Figs. 7f, h and 8d). The overall timing of the SNOW/SAC was generally more accurate and the model had a higher NSE (0.60) than the SAST/SAC (0.11) for the East basin (Table 4). Both models had the highest Pbias during 1992, the driest year on record (Fig. 8a).

The SAST had significantly higher melt rates in five of the 13 years during periods of rapid melt. Therefore, despite the later onset of melt, the SAST melted the snowpack an average of 2.5 days earlier than the SNOW17 (Table 3). 1995 is an example of where the late onset of significant melting caused early discharge events to be missed, leading to excess SWE in the spring and resulting in overestimated peak discharge events around late May and early June 1995 (Figs. 3d and 7g, h). Despite the error in streamflow timing, the volume error is small (Fig. 8e). While the SNOW/SAC had the better overall performance, simulated streamflow is similar in years where the melt pattern is also similar between the two model (e.g. 1984 and 1988; Fig. 7a–d).

Due to differences in melt pattern and timing, several SACSMA parameters values were largely different between the two snow models (Table 2). The SI value (which initiates

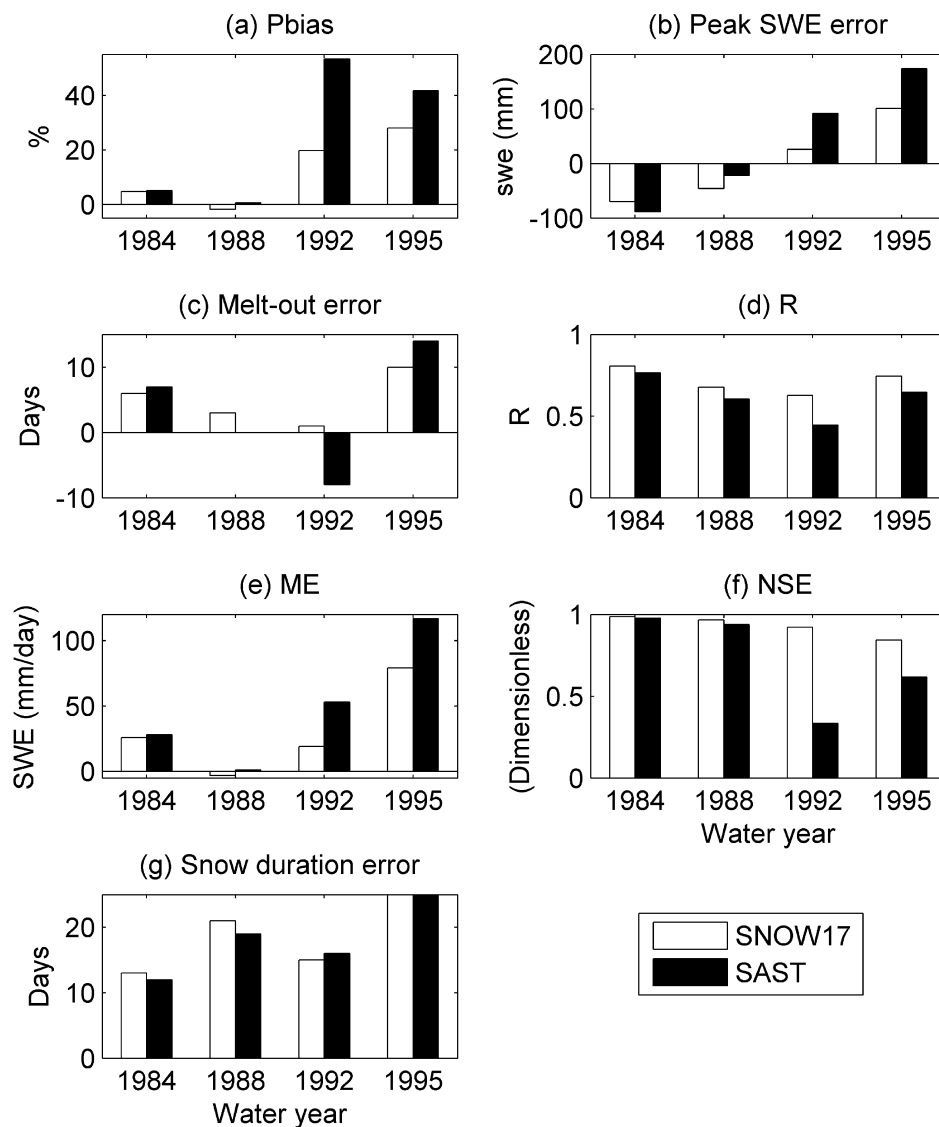


Figure 4 Mean modeled snowpack statistics for the East basin study site for water years 1984, 1988, 1992, and 1995: (a) mean daily percent bias (Pbias), (b) simulated seasonal peak SWE error, (c) difference in the timing of modeled snowpack melt in the spring compared to the observed t (melt-out error) (modeled minus observed), (d) correlation, (e) mean daily error, (f) Nash–Sutcliffe efficiency score for daily SWE, and (g) error in the number of days during the water year that the model had simulated snow compared to the observed.

Table 3 SNOW17 and SAST model simulation summary statistics for 13-year record at the RME snow pillow site

	SNOW17	SAST
NSE (dimensionless)	0.95	0.84
Pbias (%)	9.40	13.00
ME (mm/day)	26.7	34.2
R (dimensionless)	0.74	0.60
Peak SWE error (mm)	3.90	25.60
Melt-out error (days)	3.0	0.5
Snow duration error (days)	16.8	17.5

the application of the ADC) was extensively tested for the East basin because an equal optimum for both models was not found. SI was set to 50 and 200 mm for the SAST/SAC

and SNOW/SAC, respectively (Table 2). A lower SI results in longer period of complete snow cover in the SAST/SAC simulations. Due to the tendency of the SAST to melt more

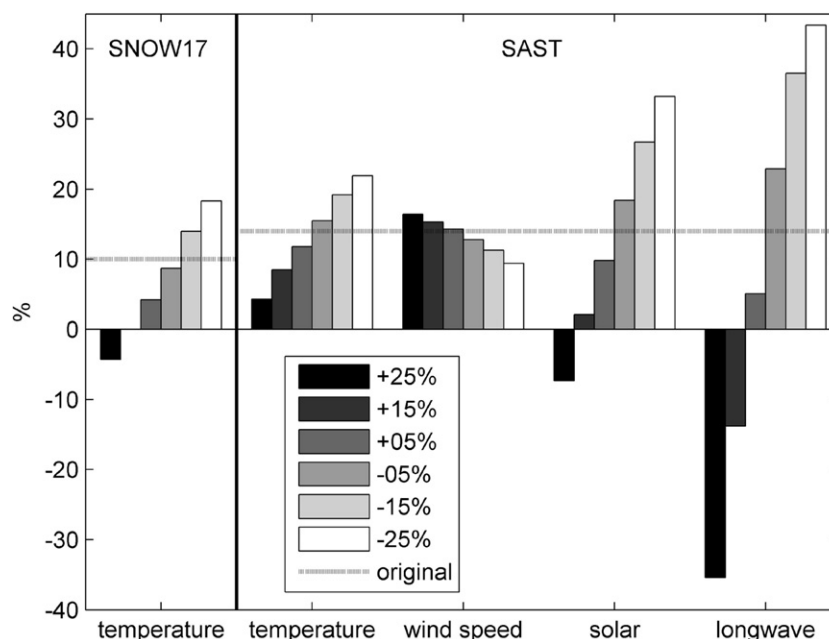


Figure 5 Mean daily percent bias (Pbias) in simulated SWE for the East basin study site for the original model run (dotted line) and model runs with error (see legend) added to the input data for the SNOW17 and the SAST. Errors in relative humidity produced less than a 1% change in percent bias for the SAST.

rapidly, the snow covered area declines rapidly once the SI value is met. In order to maintain sufficient melt water output to reproduce adequate discharges, the decrease in snow covered area had to be delayed by setting a lower SI value. The ADC functions slightly differently for the two models (in the SNOW17 the ADC also modifies the melt rate), so varying the value of SI to compensate for the varying model structure is not unreasonable.

The late melt onset, fast melt rates, and small SI in the SAST, resulted in larger lower zone storages (LZTWM and LZFSM) and higher percolation rates (ZPERC) in the SAST/SAC (Table 2) in order to move excess water quickly into the lower soil zone, dampen the large SAST melt outputs, and allow the simulated streamflow peaks to match the observed more accurately. The LZTWM can function as a sink for excess water in the system and the larger LZTWM value in the SAST/SAC is a likely cause for the 32 mm deficit in the March through June streamflow volume, compared to only a 8 mm deficit in the SNOW/SAC (Table 4).

Tollgate watershed

The 13-year average NSE values improved for the SAST/SAC in the Tollgate watershed simulations, increasing to 0.31 (Table 4). If 1989 is removed from the 13-year average, the NSE increases to 0.43 and peak discharge error is lowered to 0.34 mm/day. In 1985, 1986, and 1989 the SAST had very little melt from December through March and rapid melting in April. In these instances, the peak discharge was overestimated by the SAST/SAC (not shown). With the exception of the 1985 accumulation period, these years are climatically similar to the other years in which the SAST/SAC tended to underestimate peak discharge.

The melt patterns between the two snow models varied to a larger degree in the Tollgate (Fig. 9a, c, e, and g),

and most significantly in 1992 where the SNOW17 had significantly more melt during November–December than the SAST. The SAST displays a later onset of melt and accelerated late spring melt rates similar to what was observed for the East; however, average peak discharge errors of the SAST/SAC are lower than those of the SNOW/SAC (Table 4 and Fig. 10b). Water years 1995 (Fig. 9h) and 1996 (not shown) revealed a problem with late melt for the SAST/SAC in the East basin, but these years do not show the same simulation errors in the Tollgate basin.

Four parameters had greater than 20% difference between the SAST/SAC and SNOW/SAC (UZFWM, ZPERC, LZFSM, and KROUTE) in Tollgate (Table 2). However, the values did not tend towards the optimization bounds as was seen in the smaller East basin. In the larger Tollgate, the simulations were less sensitive to the SI and a single value of 300 mm worked well for both models. It is likely that the uncertainty resulting from using the same value in both models is masked by other problems (i.e. data averaging) introduced at the larger scale.

Discussion

Based on 13 years of simulations, the SNOW17 performed consistently better than the SAST in both the East and Tollgate basins. Much of the difference between the estimated SWE and discharge from the two snow models are linked to their respective melt patterns and rates. For those years when the SAST had minimal winter ablation and late snow-melt, large differences in estimated daily and seasonal peak SWE errors were observed between the two models. The differences were only significant in four of the 13 years in the East basin but occurred in both wet and dry years. The SAST had a rapid spring melt rate resulting in an average earlier

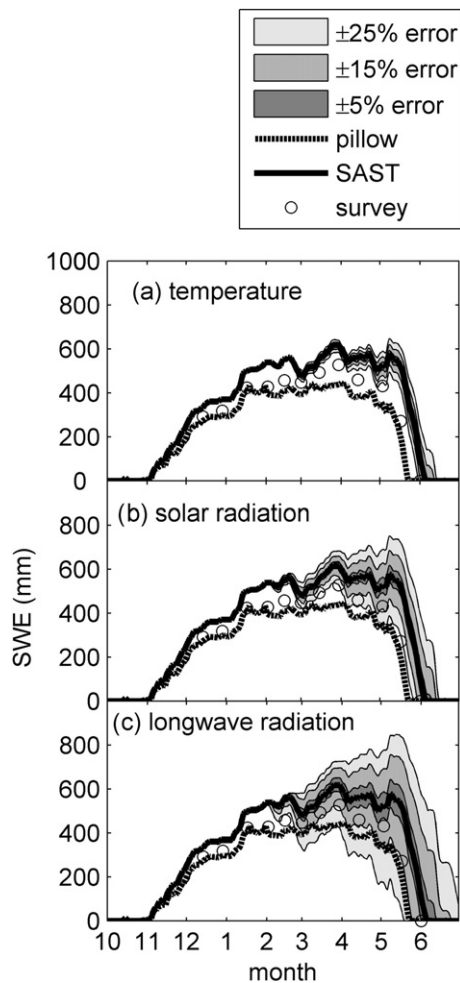


Figure 6 Impact on simulated daily SWE for the East basin for WY 1995 due to errors added to the SAST model input.

timing of complete pack melt-out as compared to the SNOW17, but resulted in overestimation of peak discharge. The SAST/SAC performance improved in the Tollgate, but on average, had lower NSE scores and larger peak discharge errors relative to the SNOW/SAC. However, for several years (1984, 1990, 1992, 1993, and 1995) the SAST/SAC had equal or higher NSE scores and lower peak discharge errors.

Point-scale snow model comparisons

The SAST tended to miss mid-winter melting episodes leading to high spring snow depths. Mid-winter ablation is a source of uncertainty in many snow schemes; in the PILPS-2(d) model comparison study, the participating land surface models were found to have markedly different early season ablation patterns (Slater et al., 2001). In the PILPS-2(e) experiment, which compared 21 land-surface schemes in several high-latitude basins, the SAST overestimated the accumulated SWE at three of the test sites (Nijssen et al., 2003). A strong link between winter over-accumulation errors and new snow albedo parameterization in the SAST was found by Xue et al. (2003); however, at the East basin the new snow albedo value was insignificant compared to the snow layer thickness parameters. The thickness of the

top two snow layers is important for reasonable simulation of diurnal temperature changes and heat conduction within the pack (Sun et al., 1999; Jin et al., 1999a). Under heavy snow conditions an excessively thick second layer leads to incorrect simulation of the ablation timing in the SAST. Reducing the thickness of the top two snow layers in the SAST through calibration significantly reduced the SWE over-accumulation in the years with the largest snowpack in the East basin. Although the hydrologic community is divided on the need for calibration of physically based models (Gupta et al., 1998; Kirchner, 2006) (such as an energy balance snow model), “effective parameters” that cannot be directly obtained from field measurements arguably require calibration (Gupta et al., 1999; Hogue et al., 2006a). Previous studies and the sensitivity of the SAST to selected parameters shown here, indicate that the use of an energy balance model would not alleviate calibration requirements for applications such as streamflow forecasting.

SNOW17 parameters calibrated to a nearby basin required little adjustment, and only the MFMAX and PLWHC were found to have significant impact on the results. MFMAX dominates melt computations in mountainous regions where snow cover builds throughout the winter and doesn’t melt until spring (Anderson, 2002); as such, this value should be adjusted to specific sites. In previous work, the uncalibrated SNOW17 produced an early onset of melt (Lundquist and Flint, 2006) and completely melted the snowpack earlier than observed (Etchevers et al., 2002). Given the highly conceptual nature of the SNOW17, studies that analyzed the model without site specific parameter calibration are difficult to contrast against the work presented here. Simulations in the RCEW show that a well calibrated SNOW17 model is highly accurate in dry to wet periods for point and watershed scale simulations.

The SAST requires more data, leading to increased opportunity for input uncertainty to be propagated through to model simulations. Data biases had the greatest effect on the SWE in the late winter and spring and likely contributed to the faster spring melt rates during this period. Previous studies have shown that longwave radiation estimates tend to be positively biased (Marks and Dozier, 1979; Fierz et al., 2003). The methodology used for estimating longwave radiation at RCEW was tested using 5 years of observed climate data at the Mammoth Mountain snow study site (Mammoth) located in east-central California (<http://neige.bren.ucsb.edu/mmsa/>); longwave radiation observations were available during the ablation seasons from 1992 to 1996. At Mammoth, our estimated longwave radiation was 11% higher than the observed, varying between +7% and +14% from season to season (Franz, 2006). The biased longwave increased the onset and rate of spring snow melt in the SAST simulations for Mammoth. Data biases may vary by location, however, a positively biased longwave input would explain, in part, the rapid melt rates observed at the RCEW sites. Given the relative model insensitivity to longwave biases in fall and early winter, the over-accumulation of SWE in the mid-winter period is more likely linked to errors in the albedo computation. Energy balance models suffer from feedbacks between errors in SWE and subsequently albedo and the radiation balance in the model (Mitchell et al., 2004), exacerbating tendencies to over- or under-accumulate the snowpack. The snow age albedo estimation methods (such

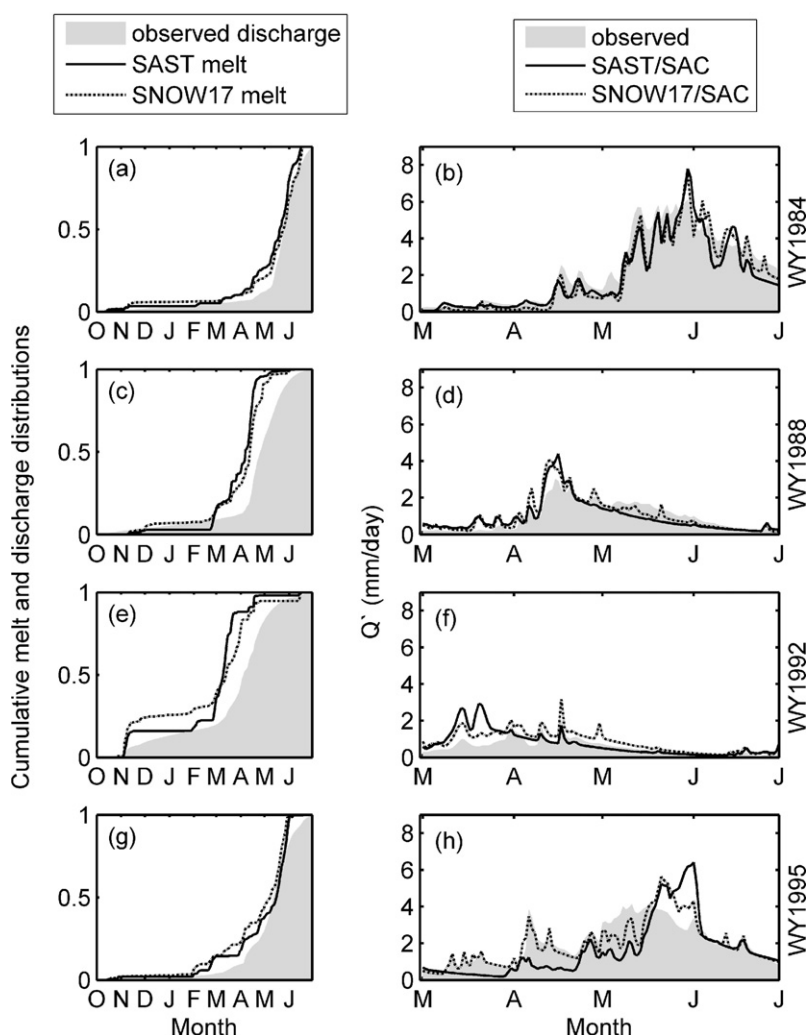


Figure 7 Cumulative distribution of simulated melt and observed discharge (a, c, e, g), and simulated and observed daily discharge (b, d, f, h) for the East basin watershed for water years (WY) 1984, 1988, 1992, and 1995.

Table 4 Model simulation summary statistics for 13-year record in the RME and TOLL watersheds for the snow models coupled to the Sacramento model

	RME		TOLL	
	SNOW17/SAC	SAST/SAC	SNOW17/SAC	SAST/SAC
NSE (dimensionless)	0.60	0.11	0.58	0.31
Pbias (%)	7.4	−0.4	−0.5	0.5
RMSE (mm/day)	0.13	0.15	0.08	0.10
Peak discharge error (mm/day)	9.4	13.0	1.8	1.2
Peak discharge Timing error (days)	3.2	−1.0	−7.3	−7.1
March 1–June 30 volume error (mm)	−8.2	−31.7	4.2	26.9

as used here) are accurate during melting periods, but not as accurate during mid-winter, non-melt periods (Etchevers et al., 2004). Alternative albedo estimation methods were tested initially and found to have little impact on overall simulated SWE (Franz, 2006), but will be revisited in future studies.

The SNOW17 was slightly more sensitive to temperature data biases than the SAST. This contrasts findings by Lei

et al. (2007) in which the SNOW17 was shown to be less sensitive to biases in temperature than an energy balance model. In both studies, however, the energy balance model was least sensitive to wind speed and most sensitive to radiation model forcings. The complex interactions between the energy balance snow model and data errors presented here support findings of Lei et al. (2007) which state that better estimates of data are needed to run an energy balance snow

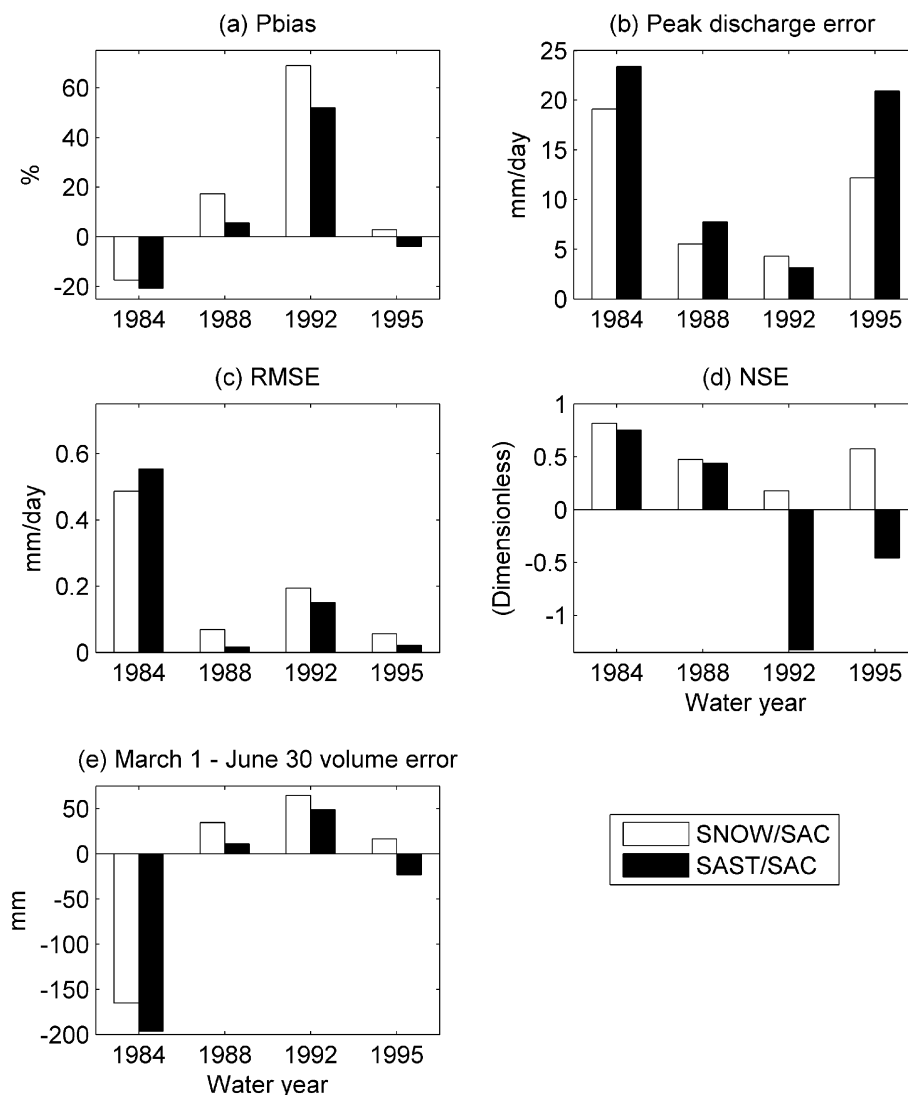


Figure 8 Mean modeled streamflow statistics by year for the East basin study for March 1st to June 30th for 1984, 1988, 1992, and 1995: (a) mean daily percent bias (Pbias), (b) peak discharge error, (c) daily root mean square error (RMSE), (d) Nash–Sutcliffe efficiency (NSE), and (e) error in total discharge.

model for operational forecasting. Accurate representation of the radiation variables to which the SAST is most sensitive will be critical for use of the model in hydrologic prediction.

The snow models have only temperature and precipitation as common inputs, which determine the accumulation of snow. A study conducted in the Sheep Creek basin of the Reynolds Creek Watershed pointed to model errors as the reason for overestimated peak accumulation produced by the Utah Energy Balance model (in both distributed and lumped mode) (Luce et al., 1999). However, the common trend shown here by both the SAST and SNOW17 to overaccumulate early in the snow season indicates that errors in temperature, precipitation and/or observed SWE are contributing to the uncertainty in model predictions. In addition, the rain/snow cutoff parameter value of 1 °C may require adjustment in some years, such as 1992, 1993, 1995, and 1996, where the modeled SWE was overestimated by both the SNOW17 and SAST. Varying parameters on an annual basis is not typically considered in hydrologic models,

however, adjusting the rain/snow cutoff parameter for climatic conditions and identifying the interaction between parameters and sources of uncertainty will be explored in future studies.

Watershed scale comparisons

An areal depletion curve (ADC) was used in place of a distributed snow model application to account for the influence of terrain and snow redistribution in the Reynolds creek watershed. Marks and Winstral (2001) showed that runoff generated using the SNOBAL, an energy balance snow model, at both the ridge and the snow pillow sites of the East basin under-represented discharge from the basin because relatively more water was contributed to discharge due to large drifts within the watershed. Mean wind speed was significantly higher during 1993, 1995, and 1996, coinciding with the greatest simulated SWE overestimations by the snow models. Drifting would be most likely in these

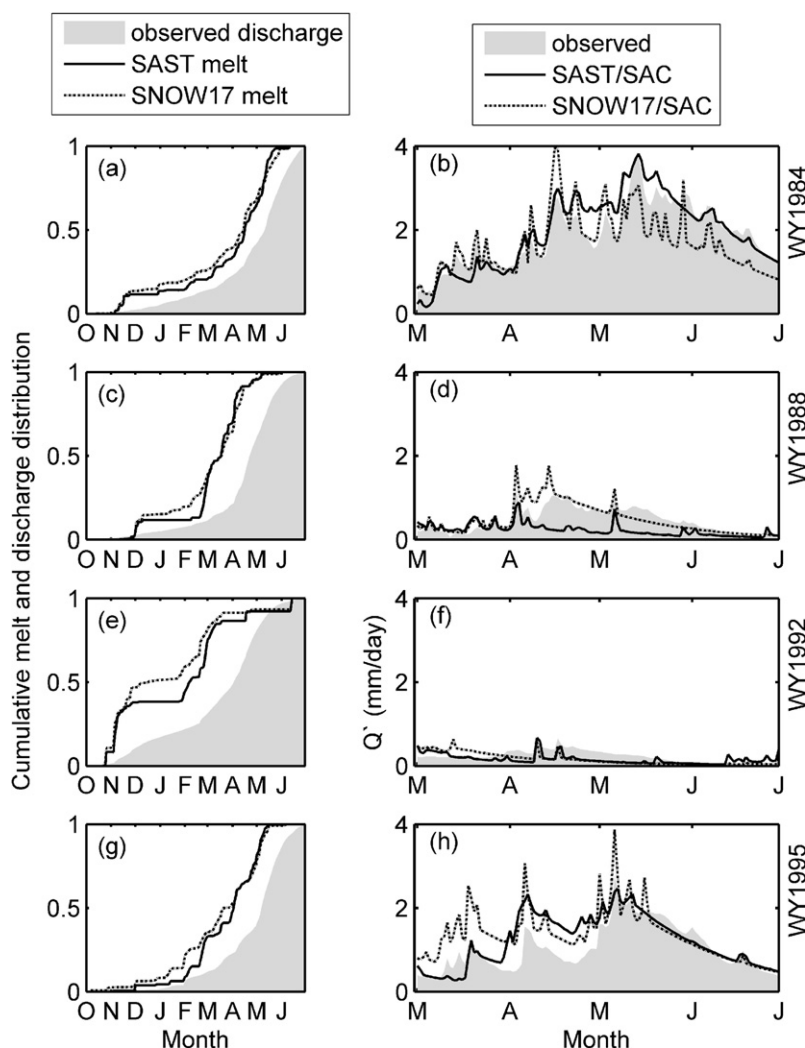


Figure 9 Cumulative distribution of simulated melt and observed discharge (a, c, e, g), and simulated and observed daily discharge (b, d, f, h) for the Tollgate basin for water years 1984, 1988, 1992, and 1995.

years. However, no relationship between the errors in peak discharge or streamflow volume with high mean wind speed is obvious, indicating that the drifting did not have a dominant impact on simulated streamflow errors or the ADC was able to account for any effects from drifting. The ADC method appeared better suited to the larger watershed based on SAST/SAC discharge patterns. The potential use of a lumped energy balance model for snow simulation when detailed subbasin conditions are not necessary has also been shown by Luce et al. (1999), who found that lumped snow model simulations using an areal depletion curve agreed with distributed versions of the same model.

SAC/SAST parameters calibrated within reasonable ranges identified by NWS guidance and other studies, indicating compatibility between the SAST and SACSMA. Differences in the calibrated SAST/SAC and SNOW/SAC parameters arose because (1) the diurnal and seasonal melt patterns from the snow models are quite different, (2) the SACSMA is a conceptual model, therefore the parameters are not direct representatives of basin characteristics and must be calibrated with specific forcings, and (3) parameter interaction can result in multiple parameter sets which may

have similar performance. The calibration and application of the SACSMA confirmed that model parameters are sensitive to characteristics in the input and that the SACSMA requires calibration specific to the data sources in addition to the watershed (Finnerty et al., 1997; Yilmaz et al., 2005).

Concluding remarks

This study was undertaken to determine the feasibility of using a common energy balance model in lieu of the current NWS snow model for streamflow forecasting. We emphasize that the SAST was used as a proxy for the class of energy balance models and this study was not meant to target the performance or potential deficiencies of the SAST. The SAST estimated point- and basin-scale processes as accurately as the SNOW17 for most years even with simple estimation of basin-average climate forcing and longwave radiation. Our conclusions do not preclude the use of an energy balance model in operational forecasting, however, relatively large uncertainty still exists in the predictive skill of the energy balance model relative to existing procedures. Several challenges remain before the application of an en-

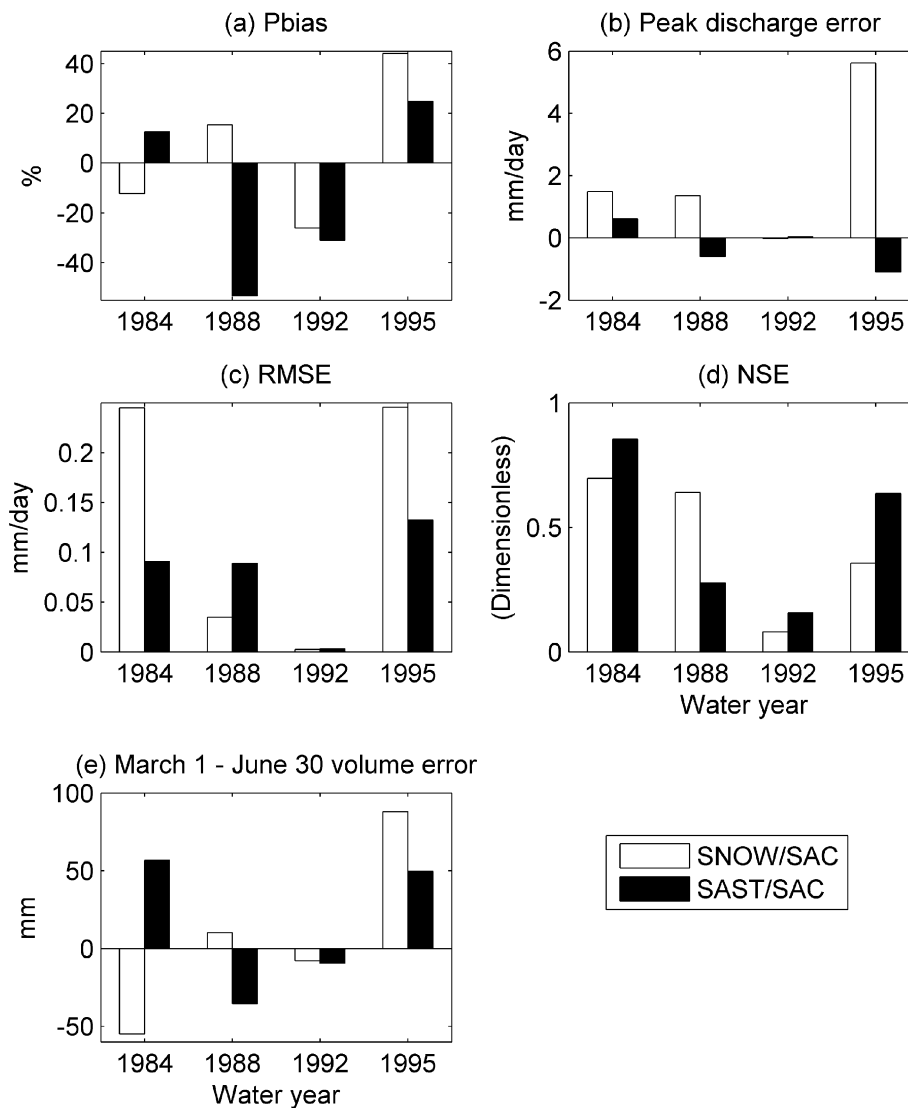


Figure 10 Mean modeled streamflow statistics by year for the Tollgate study for March 1st to June 30th for 1984, 1988, 1992, and 1995: (a) mean daily percent bias (Pbias), (b) peak discharge error, (c) daily root mean square error (RMSE), (d) Nash–Sutcliffe efficiency (NSE), and (e) error in total discharge.

ergy balance model for operational predictions can be realized:

- The SAST performance quality will be influenced to a greater extent by biased data than the SNOW17 model. Data error estimation and bias correction will be challenging due to snow-energy balance feedbacks within the model and difficulty in estimating biases for multiple data streams.
- The need to calibrate the SAST will not alleviate existing calibration requirements. The ease of the SNOW17 calibration and ease to which parameters were transferred from an operational basin to the research basins can be attributed to the long history of use of this model by the NWS. An understanding of parameter ranges and areal depletion curves is well documented for the SNOW17. A comparable understanding of parameter ranges and sensitivities will be required for an energy balance model.
- The difference in SACSMA model parameters illustrates that a new snow model will require extensive recalibration of the SACSMA, substantial investigation will be needed to optimally calibrate both an energy balance model and the SACMSA model and to understand the associated biases and uncertainty.
- Current analysis of energy balance models is limited due to inadequate basin-scale hydrologic observations. Long-term analysis of the energy balance model in various climates and locations using data sources similar to that which will be used operationally will be needed.

Historical model simulation analysis is an important first step in model evaluation; however, it does not necessarily provide information about how a model will perform under forecasting conditions. In a follow-on paper, hindcasting techniques are added to our model analysis to quantify the snow model skill for ensemble streamflow prediction. Through these series of papers we set forth a framework

through which alternatives snow models may be evaluated against current operational models.

Acknowledgments

Primary financial support provided by the NASA GEWEX Program, Grant Number NNG04GM24G, and NASA EOS, Grant Number NAG5-11044 is appreciated. Financial support provided by the University of California, Irvine Newkirk Center for Science and Society is also greatly appreciated. Thanks to Jiming Jin for providing technical support for the SAST model and to several NWS personnel, Victor Koren, Harold Opitz, and Joe Intermill, for providing model guidance and data. Data for Reynolds Creek Watershed was provided by the US Department of Agriculture Northwest Watershed Research Center. We would also like to thank the University of California, Irvine for support of this work.

References

- Anderson, E.A., 1968. Development and testing of snow pack energy balance equations. *Water Resour. Res.* 4 (1), 19–37.
- Anderson, E.A., 1973. National Weather Service River Forecast System-Snow Accumulation and Ablation Model. NOAA Technical Memorandum: NWS Hydro-17, US National Weather Service.
- Anderson, E.A., 1976. A point energy and mass balance model of a snow cover. NOAA Technical Report NWS 19, US National Weather Service, 150 pp.
- Anderson, E.A., 2002. Calibration of conceptual hydrologic models for use in river forecasting. Office of Hydrologic Development, US National Weather Service, Silver Spring, MD.
- Bowling, L.C., Lettenmaier, D.P., Nijssen, B., Graham, L.P., Clark, D.B., Maayar, M.E., Essery, R., Goers, S., Gusev, Y.G., Habets, F., van den Hurk, B., Jin, J., Kahan, D., Lohmann, D., Ma, X., Mahanama, S., Mocko, D., Nasonova, O., Niu, G.Y., Samuelsson, P., Shmakin, A.B., Takata, K., Verseghy, D., Viterbo, P., Xia, Y., Xue, Y., Yang, Z.-L., 2003. Simulation of high-latitude hydrological processes in the Torne-Kalix basin. PILPS Phase 2(e) 1: Experiment description and summary intercomparisons. *Global Planet. Change* 38, 1–30.
- Boyle, D.P., Gupta, H.V., Sorooshian, S., 2000. Toward improved calibration of hydrologic models: combining the strengths of manual and automatic methods. *Water Resour. Res.* 36 (12), 3663–3674.
- Brubaker, K., Rango, A., Kustas, W., 1996. Incorporating radiation inputs into the snowmelt runoff model. *Hydrol. Process.* 10, 1329–1343.
- Burnash, R.J., Ferral, R.L., McGuire, R.A., 1973. A Generalized Streamflow Simulation System Conceptual: Modeling for Digital Computers. Joint Federal-State River Forecast Center, Sacramento, CA.
- Carroll, T., Cline, D., Fall, G., Nilsson, A., Li, L., Rost, A., 2001. NOHRSC operations and the simulation of snow cover properties for the conterminous US. In: *Proceedings of the 69th Annual Western Snow Conference*, April 2001, Sun Valley, ID.
- Crawford, T.M., Duchon, C.E., 1999. An improved parameterization for estimating effective atmospheric emissivity for use in calculating daytime downwelling longwave radiation. *J. Appl. Meteorol.* 38, 474–480.
- Crow, W.T., Entekhabi, D., Koster, R.D., 2006. Multiple spaceborne water cycle observations would aid modeling. *EOS* 87 (15), 149–153.
- Day, G.N., Schaake, J.C., Ellis, J.H., 1989. A direction toward improved streamflow forecasting in the western mountains. In: *Proceedings of the 57th Annual Western Snow Conference*, Fort Collins, CO, pp. 79–89.
- Dickinson, R.E., Henderson-Sellers, A., Kennedy, P.J., 1993. Biosphere–Atmosphere Transfer Scheme (BATS) Version 1e as Coupled to the NCAR Community Climate Model. NCAR Technical Note NCAR/TN-387+STR, Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, CO.
- Dingman, S.L., 2002. *Physical Hydrology*, second ed. Prentice Hall, New Jersey, p. 646.
- Duan, Q., Gupta, V.K., Sorooshian, S., 1992. Effective and efficient global optimization for conceptual rainfall–runoff models. *Water Resour. Res.* 28, 1015–1031.
- Essery, R., Martin, E., Douville, H., Fernandez, A., Brun, E., 1999. A comparison of four snow models using observations from an alpine site. *Clim. Dynam.* 15, 583–593.
- Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, N.O., Pyles, R.D., Schlosser, A., Shmakin, A.B., Smirnova, T.G., Strasser, U., Verseghy, D., Yamazaki, T., Yang, Z.-L., 2002. SnowMIP, an intercomparison of snow models first results. In: *Proceedings of the International Snow Science Workshop*.
- Etchevers, P., Martin, E., Brown, R., Fierz, C., Lejeune, Y., Bazile, E., Boone, A., Dai, Y.J., Essery, R., Fernandez, A., Gusev, Y., Jordan, R., Koren, V., Kowalczyk, E., Nasonova, N.O., Pyles, R.D., Schlosser, A., Shmakin, A.B., Smirnova, T.G., Strasser, U., Verseghy, D., Yamazaki, T., Yang, Z.-L., 2004. Validation of the energy budget of an alpine snowpack simulated by several snow models (SnowMIP project). *Ann. Glaciol.* 38, 150–158.
- Fierz, C., Riber, P., Adams, E.E., Curran, A.R., Föhn, P.M.B., Lehning, M., Plüss, C., 2003. Evaluation of snow-surface energy balance models in alpine terrain. *J. Hydrol.* 282, 76–94.
- Finnerty, B.D., Smith, M.B., Seo, D.J., Koren, V., Moglen, G.E., 1997. Space-time scale sensitivity of the Sacramento model to radar-gage precipitation inputs. *J. Hydrol.* 203 (1–4), 21–38.
- Franz, K.J., 2006. Characterization of the comparative skill of conceptual and physically-based snow models for streamflow prediction. Ph.D. Dissertation, University of California, Irvine.
- Gupta, H.V., Bastidas, L.A., Sorooshian, S., Shuttleworth, W.J., Yang, Z.-L., 1999. Parameter estimation of a land surface scheme using multicriteria methods. *J. Geophys. Res.* 104 (D16), 19491–19503.
- Gupta, H.V., Sorooshian, S., Yapo, P.O., 1998. Toward improved calibration of hydrologic models: multiple and noncommensurable measures of information. *Water Resour. Res.* 34 (4), 751–763.
- Hanson, C.L., 2001. Long-term precipitation database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resour. Res.* 37 (11), 2831–2834.
- Hanson, C.L., Marks, D., van Vactor, S.S., 2001. Long-term climate database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resour. Res.* 37 (11), 2839–2841.
- Hogue, T.S., Bastidas, L.A., Gupta, H.V., Sorooshian, S., 2006a. Evaluating model performance and parameter behavior for varying levels of land surface model complexity. *Water Resour. Res.* 42 (8), W08430. doi:10.1029/2005WR00444.
- Hogue, T.S., Gupta, H.V., Sorooshian, S., 2006b. A “User-Friendly” approach to parameter estimation in hydrologic models. *J. Hydrol.* 320, 202–217.
- Hogue, T.S., Sorooshian, S., Gupta, H.V., Holz, A., Braatz, D., 2000. A multistep automatic calibration scheme for river forecasting models. *J. Hydrometeorol.* 1, 524–542.
- Jin, J., Gao, X., Yang, Z.-L., Bales, R.C., Sorooshian, S., Dickinson, R.E., Sun, S.F., Wu, G.X., 1999a. Comparative analyses of physically based snowmelt models for climate simulations. *J. Climate* 12, 2643–2657.
- Jin, J., Gao, X., Sorooshian, S., Yang, Z.-L., Bales, R., Dickinson, R.E., Sun, S.F., Wu, G.X., 1999b. One-dimensional snow water

- and energy balance mode for vegetated surfaces. *Hydrol. Process.* 13, 2467–2482.
- Jordan, R., 1991. A one-dimensional temperature model for a snow cover. US Army Corps of Engineers, Cold Regions Research and Engineering Laboratory, Special Report 91-16, p. 49.
- Kimball, B.A., Idso, S.B., Aase, J.K., 1982. A model of thermal radiation from partly cloudy and overcast skies. *Water Resour. Res.* 18 (4), 931–936.
- Kirchner, J.W., 2006. Getting the right answers for the right reasons: linking measurements, analyses, and models to advance the science of hydrology. *Water Resour. Res.* 42, W03S04. doi:10.1029/2005WR00436.
- Koren, V.I., Smith, M., Wang, D., Zhang, Z., 2000. Use of soil property data in the derivation of conceptual rainfall–runoff model parameters. In: Paper 2.16, American Meteorology Society, January 9–14, 2000, Long Beach, CA.
- Lang, H., 1986. Forecasting meltwater runoff from snow-covered areas and glacier basins. In: Kraijenhoffand, D.A., Moll, J.R., Reidel, D. (Eds.), *River Flow Modeling and Forecasting*. Dordrecht, Holland, pp. 99–127.
- Lang, H., Braun, L., 1990. On the information content of air temperature in the context of snow melt estimation. *Hydrology of Mountainous Areas*. IAHS Publication No. 190. L. Molnar, pp. 347–354.
- Lehning, M., Völsch, I., Gustafsson, D., Nguyen, T.A., Stähli, M., Zappa, M., 2006. ALPINE3D: a detailed model of mountain surface processes and its application to snow hydrology. *Hydrol. Process.* 20, 2111–2128.
- Lei, F., Koren, V., Smith, M., Moreda, F., 2007. A sensitivity study of an energy budget snow accumulation and ablation model. In: Paper J.6.4, American Meteorological Society, San Antonio, TX.
- Luce, C.H., Tarboton, D.G., Cooley, K.R., 1999. Sub-grid parameterization of snow distribution for an energy and mass balance snow cover model. *Hydrol. Process.* 13, 1921–1933.
- Lundquist, J.D., Flint, A.L., 2006. Onset of snowmelt and streamflow in 2004 in the western United States: how shading may affect spring streamflow timing in a warmer world. *J. Hydrometeorol.* 7, 1199–1217.
- Marks, D., Dozier, J., 1979. A clear-sky longwave radiation models for remote alpine areas. *Arch. Met. Geoph. Biokl. Ser. B* 27, 159–187.
- Marks, D., Winstral, A., 2001. Comparison of snow deposition, the snow cover energy balance, and snowmelt at two sites in a semiarid mountain basin. *J. Hydrometeorol.* 2 (3), 213–227.
- Marks, D., Domingo, J., Susong, D., Link, T., Garen, D., 1999. A spatially distributed energy balance snowmelt model for application in mountain basins. *Hydrol. Process.* 13, 1935–1959.
- Marks, D., Cooley, K.R., Robertson, D.C., Winstral, A., 2001. Long-term snow database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resour. Res.* 37 (11), 2835–2838.
- Maurer, E.P., 2007. Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Climatic Change* 82, 309–325.
- McEnery, J., Ingram, J., Duan, Q., Adams, T., Anderson, L., 2005. NOAA's advanced hydrologic prediction service: building pathways for better science in water forecasting. *Bull. Am. Meteor. Soc.* 86 (3), 375–385.
- Mitchell, K.E., Lohmann, D., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Cosgrove, B.A., Sheffield, J., Duan, Q.Y., Luo, L.F., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Lettenmaier, D.P., Marshall, C.H., Entin, J.K., Pan, M., Shi, W., Koren, V., Meng, J., Ramsay, B.H., Bailey, A.A., 2004. The multi-institution North American Land Data Assimilation System (NLDAS): utilizing multiple GCIIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.: Atmos.* 109 (D7) article no. D07590. doi:10.1029/2003JD003823.
- Mote, P.W., 2003. Trends in snow water equivalent in the Pacific Northwest and their climate causes. *Geophys. Res. Lett.* 30 (12), 1601. doi:10.1029/2003GL01725.
- Mote, P.W., Hamlet, A.F., Clark, M.P., Lettenmaier, D.P., 2005. Declining mountain snowpacks in western north America. *Bull. Am. Meteor. Soc.* 86 (1), 39–49.
- Nijssen, B., Bowling, L.C., Lettenmaier, D.P., Clark, D.B., Maayar, M.E., Essery, R., Goers, S., Gusev, Y.G., Habets, F., van den Hurk, B., Jin, J., Kahan, D., Lohmann, D., Ma, X., Mahanama, S., Mocko, D., Nasonova, O., Niu, G.Y., Samuelsson, P., Shmakin, A.B., Takata, K., Verseghy, D., Viterbo, P., Xia, Y., Xue, Y., Yang, Z.-L., 2003. Simulation of high-latitude hydrological processes in the Torne-Kalix basin. PILPS Phase 2(e) 2: Comparison of model results with observations. *Global Planet. Change* 38, 31–53.
- NRC, 2006. Towards Advanced Hydrologic Prediction Service (AHPs). National Research Council Committee to Assess the National Weather Service Advanced Hydrologic Prediction Service Initiative, S. Sorooshian chair, National Academies Press, Washington, DC, p. 74.
- NWS, 2004. The National Weather Service River Forecast System User's Manual. <<http://www.nws.noaa.gov/ohd/hrl/>> (referenced version was downloaded on January 9, 2004).
- Ohmura, A., 2001. Physical basis for the temperature-based melt-index method. *J. Appl. Meteorol.* 40, 753–761.
- Perkins, T.R., 1988. Seasonal streamflow forecasting in the upper Rio Grande basin by incorporating the use of SNOTEL data in the SSARR hydrologic model. In: Proceedings of the 56th Annual Western Snow Conference, Kalispell, MT, pp. 58–69.
- Pierson, F.B., Slaughter, C.W., Cram, Z.K., 2001. Long-term stream discharge and suspended-sediment database, Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resour. Res.* 37 (11), 2857–2861.
- Schaake, J., Franz, K., Bradley, A., Buizza, R., 2006. The Hydrologic Ensemble Prediction Experiment (HEPEX). *Hydrol. Earth Syst. Discuss.* 3, 3321–3332.
- Shamir, E., Georgakakos, K.P., 2006. Distributed snow accumulation and ablation modeling in the North American River basin. *Adv. Water Resour.* 29, 558–570.
- Slater, A.G., Schlosser, C.A., Desborough, C.E., Pitman, A.J., Henderson-Sellers, A., Robock, A., Vinnikov, K.Ya., Mitchell, K., Boone, A., Braden, H., Chen, F., Cox, P.M., De Rosnay, P., Dickinson, R.E., Dai, Y.-J., Duan, Q., Entin, J., Etchevers, P., Gedney, N., Gusev, Y.E., Habets, F., Kim, J., Koren, V., Kowalczyk, E.A., Nasonova, O.N., Noilhan, J., Schaake, S., Shmakin, A.B., Smirnova, T.G., Verseghy, D., Wetzel, P., Xue, Y., Yang, Z.-L., Zeng, Q., 2001. The representation of snow in land surface schemes: results from PHILPS 2(d). *J. Hydrometeorol.* 2, 7–25.
- Slaughter, C.W., Marks, D., Flerchinger, G.N., Van Vactor, S.S., Burgess, M., 2001. Thirty-five years of research data collection at the Reynolds Creek Experimental Watershed, Idaho, United States. *Water Resour. Res.* 37 (11), 2819–2823.
- Sorooshian, S., Duan, Q., Gupta, V.K., 1993. Calibration of rainfall–runoff models: application of global optimization to the Sacramento Soil Moisture Accounting model. *Water Resour. Res.* 29, 1185–1194.
- Sorteberg, H.K., Engeset, R.V., Udnæs, H.C., 2001. A national network for snow monitoring in Norway: snow pillow verification using observations and models. *Phys. Chem. Earth* 26 (10–12), 723–729.
- Steiner, A., 2001. Effects of land cover change on regional atmospheric chemistry and climate in China. Interim Report IR-01-017, International Institute for Applied Systems Analysis, Laxenburg, Austria.
- Stewart, I.T., Cayan, D.R., Dettinger, M.D., 2004. Changes in snowmelt runoff timing in western North America under a 'business as usual' climate change scenario. *Climate Change* 62, 217–232.
- Strasser, U., Etchevers, P., Lejeune, Y., 2002. Inter-comparison of two snow models with different complexity using data from an alpine site. *Nord. Hydrol.* 33 (1), 15–26.

- Sun, W.-Y., Chern, J.-D., 2005. Validation of a one-dimensional snow–land surface model at the Sleepers River watershed. *Bound. Lay. Meteorol.* 116, 95–115.
- Sun, S., Jin, J., Xue, Y., 1999. A simple snow–atmosphere–soil transfer model. *J. Geophys. Res.* 104 (D16), 19587–19597.
- Swamy, A.N., Brovio, P.A., 1997. Modeling runoff using optical satellite remote sensing data in a high mountainous alpine catchment of Italy. *Hydrol. Process.* 11, 1475–1491.
- Tarboton, D.G., Luce, C.H., 1996. Utah Energy Balance Snow Accumulation and Melt Model (UEB). Utah Water Research Laboratory, Utah State University and USDA Forest Service, Intermountain Research Station.
- Walter, M.T., Brooks, E.S., McCool, D.K., King, L.G., Molnau, M., Boll, J., 2005. Process-based snowmelt modeling: does it require more input data than temperature-index modeling? *J. Hydrol.* 300, 65–75.
- WMO, 1986. Intercomparison of models of snowmelt runoff. Operational Hydrology Report No. 23, World Meteorological Organization, WMO-No. 646, p. 430.
- Xue, Y., Sun, S., Kahan, D.S., Jiao, Y., 2003. Impact of parameterizations in snow physics and interface processes on the simulation of snow cover and runoff at several cold region sites. *J. Geophys. Res.* 108 (D22). doi:10.1029/2002JD00317.
- Yang, Z.-L., Niu, G.-Y., 2003. The Versatile Integrator of Surface and Atmospheric processes (VISA): Part 1. *Global Planet. Change* 38, 175–189.
- Yapo, P.O., Gupta, H.V., Sorooshian, S., 1996. Automatic calibration of conceptual rainfall–runoff models: sensitivity to calibration data. *J. Hydrol.* 181, 23–48.
- Yilmaz, K.K., Hogue, T.S., Hsu, K.L., Sorooshian, S., Gupta, H.V., Wagener, T., 2005. Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *J. Hydrometeorol.* 6 (4), 497–517.
- Zappa, M., Pos, F., Strasser, U., Warmerdam, P., Gurtz, J., 2003. Seasonal water balance of an Alpine catchment as evaluated by different methods for spatially distributed snowmelt modelling. *Nord. Hydrol.* 34 (3), 179–202.